



Using neural networks and remote sensing for spatio-temporal prediction of air pollution during the COVID-19 pandemic

Marzieh Mokarram¹ · Farideh Taripناه² · Tam Minh Pham^{3,4}

Received: 27 May 2023 / Accepted: 31 October 2023 / Published online: 18 November 2023
© The Author(s), under exclusive licence to Springer-Verlag GmbH Germany, part of Springer Nature 2023

Abstract

The study aims to monitor air pollution in Iranian metropolises using remote sensing, specifically focusing on pollutants such as O₃, CH₄, NO₂, CO₂, SO₂, CO, and suspended particles (aerosols) in 2001 and 2019. Sentinel 5 satellite images are utilized to prepare maps of each pollutant. The relationship between these pollutants and land surface temperature (LST) is determined using linear regression analysis. Additionally, the study estimates air pollution levels in 2040 using Markov and Cellular Automata (CA)-Markov chains. Furthermore, three neural network models, namely multilayer perceptron (MLP), radial basis function (RBF), and long short-term memory (LSTM), are employed for predicting contamination levels. The results of the research indicate an increase in pollution levels from 2010 to 2019. It is observed that temperature has a strong correlation with contamination levels ($R^2=0.87$). The neural network models, particularly RBF and LSTM, demonstrate higher accuracy in predicting pollution with an R^2 value of 0.90. The findings highlight the significance of managing industrial towns to minimize pollution, as these areas exhibit both high pollution levels and temperatures. So, the study emphasizes the importance of monitoring air pollution and its correlation with temperature. Remote sensing techniques and advanced prediction models can provide valuable insights for effective pollution management and decision-making processes.

Keywords Air pollution · Temperature · Markov and CA-Markov chain · Neural network, industrial towns · COVID-19

Introduction

Air pollution significantly contributes to global climate change. According to the World Health Organization (WHO), air pollution causes approximately 3 million deaths worldwide each year (Guo et al. 2021; Kalajdjieski et al. 2020), with over half of these fatalities occurring in

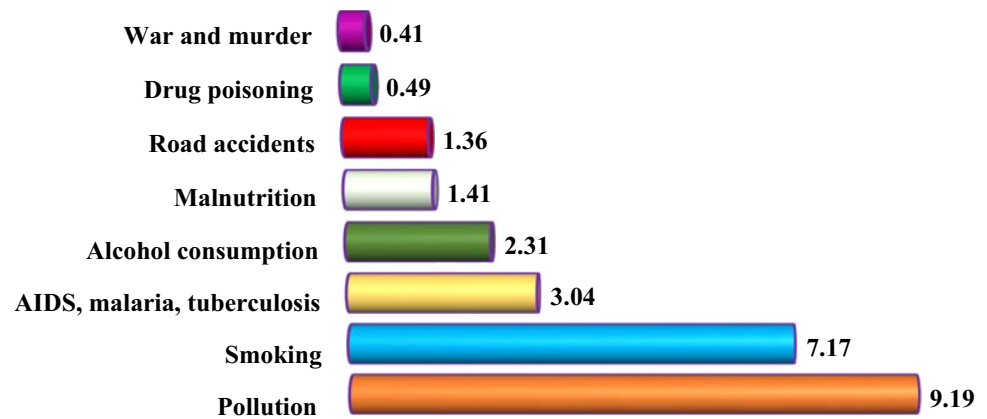
developing countries. Research indicates that the risks associated with pollution and premature death are high not only in developing nations but also in developed ones (Alimisis et al. 2018; Bauer et al. 2019). Figure 1 demonstrates that the mortality rate resulting from pollution surpasses that of other hazards and diseases globally. Many urban areas in industrialized countries are witnessing escalating pollution levels due to factors such as population growth, increased vehicle usage, and industrialization accompanied by heightened energy demand (Liu et al. 2023; Yu et al. 2021; Pedruzzi et al. 2019; Selvam et al. 2020). Therefore, it is crucial to accurately measure air pollutants with high spatial and temporal resolution at all levels to comprehend their distribution and impact and provide effective solutions to local, national, and international policymakers (Li et al. 2019b). Recently, there has been a significant rise in atmospheric pollutant levels, adversely affecting air quality, the environment, and human health. Some of the most critical air pollutants include nitrogen oxides (NO_x), sulfur dioxide (SO₂), carbon dioxide (CO₂), carbon monoxide (CO), methane (CH₄), volatile organic compounds, chlorofluorocarbons (CFCs), and aerosols (He et al. 2022).

Responsible Editor: Marcus Schulz

✉ Tam Minh Pham
phamminhtam1989@gmail.com

- ¹ Department of Geography, Faculty of Economics, Management and Social Sciences, Shiraz University, Shiraz, Iran
- ² Department of Desert Control and Management, University of Kashan, Kashan, Iran
- ³ Research Group On “Fuzzy Set Theory and Optimal Decision-Making Model in Economics and Management”, Vietnam National University, Hanoi, 144 Xuan Thuy Str., Hanoi 100000, Vietnam
- ⁴ VNU School of Interdisciplinary Studies, Vietnam National University, Hanoi, 144 Xuan Thuy Str., Hanoi 100000, Vietnam

Fig. 1 Death rates from various hazards in the world in the current period



Oil, coal, and other impure fuels contain sulfur and various organic compounds that contribute to air pollution (Jelonek et al. 2020). Additionally, air pollution arises from forest fires, soil fires, and vegetation fires, which release relatively small amounts of sulfur (Reddington et al. 2021). Notably, coal-fired power plants stand as the world's largest sources of sulfur dioxide, resulting in smoke, acid rain, and respiratory illnesses (Zhang et al. 2023; Dong et al. 2019). The significance of nitrogen oxides as pollutants and toxic gases cannot be underestimated (Czech et al. 2020). Human activities, particularly consumption, generate millions of tons of nitrogen dioxide and nitrogen oxide annually. In humid air, nitrogen dioxide leads to the production of nitric acid, causing severe metal corrosion and reduced visibility (Kwak et al. 2020). These gases also impose adverse effects on the human respiratory system and hinder plant growth (Almetwally et al. 2020; Jiang et al. 2018).

Carbon monoxide is one of the most prevalent and highly toxic air pollutants. Approximately two-thirds of carbon monoxide emissions stem from human activities, making it a significant hazard (Al-Ghussain 2019; Saevarsdottir et al. 2019). This gas is primarily generated through the incomplete combustion of carbon. Its production is not limited to burning crop residues, fossil fuels, and oxidizing CH_4 (Hu and Rein 2022). Another notable air pollutant is suspended particles or aerosols, which serve as crucial sources of air pollution with both short-term and long-term effects on human health. The impacts of these pollutants on human health include the development of cardiovascular, respiratory, and dermatological conditions that can result in premature mortality (Mor and Ravindra 2023). The presence of aerosols in the atmosphere is considered another form of air pollution, making worldwide particulate matter monitoring essential (Guo et al. 2020 and 2023a, b). Numerous environmental protection agencies employ ground stations for continuous surveillance in this regard. Suspended particles also scatter or absorb solar radiation, contributing to climate change (Tiwari and Kumar 2020; Tiwari et al. 2018).

Considering the impact of air pollution on human health, it is crucial to accurately predict and measure air pollution levels (Dong et al. 2022). Various methods exist for monitoring pollution and predicting air quality, with remote sensing being particularly significant due to its capacity to provide continuous data across time and space (Berman and Ebisu 2020; Grainger and Schreiber 2019; Vadrevu and Lasko 2018). Remote sensing relies on the transmission of air pollutant information through electromagnetic radiation, enabling the collection of data at both temporal and spatial resolutions, as well as at a vertical profile level.

Research has demonstrated that remote sensing techniques can effectively predict air quality (Selvam et al. 2020). Yuan et al. (2019) employed remote sensing data to establish that urban traffic and population growth are the primary contributors to atmospheric pollution in densely populated regions of China. Similarly, Yuan et al. (2019) reported a significant decrease (43%) in NO_2 pollution during the Asian quarantine period in 2020, attributing it to reduced industrial activities amidst the Coronavirus disease (COVID-19) era. Utilizing Moderate Resolution Imaging Spectroradiometer (MODIS) satellite images, Nichol et al. (2020) concluded that industrial activity predominantly accounts for the elevated presence of aerosols in the air.

Forecasting air pollution is vital for effective pollution management. Neural networks offer a valuable method for predicting contamination levels. Boznar et al. (1993) were pioneers in employing a neural network to forecast sulfur dioxide concentrations in polluted industrial areas of Slovenia. Several studies have investigated and utilized artificial neural networks to predict air pollution, including the works of Awan et al. (2020); Cabaneros et al. (2019). Wen et al. (2019) employed Long Short-Term Memory (LSTM) neural networks to analyze aerosol levels in the air, demonstrating the method's ability to accurately detect air pollution caused by aerosols. In a study conducted in western Iran, Maleki et al. (2019) employed neural networks to investigate Ozone (O_3), Nitrogen dioxide (NO_2), PM_{10} , $\text{PM}_{2.5}$, SO_2 , and CO

levels in the atmosphere, highlighting the potential for predictive air pollution management through Artificial neural networks (ANNs)-based models. Liu et al. (2019) also found LSTM neural networks to be highly accurate in predicting air pollution levels. Muthukumar et al. (2021) examined the spatial patterns of Particulate Matter 2.5 (PM 2.5) in Los Angeles using Graph Convolutional Networks (GCN) and LSTM, and their results further underscored the capacity of remote sensing and LSTM methods to predict air pollution with high precision.

In Iran's major cities, air pollution has emerged as a significant and palpable environmental issue (Mirsanjari et al. 2020). The presence of elevated air pollution in Iran has been observed to lead to an increase of "up to 60%" in respiratory ailments among urban residents (Santamouris and Osmond 2020). The exacerbation of cardiovascular and pulmonary diseases can be attributed to rising levels of pollutants, including NO₂, CO₂, SO₂, CO, and suspended particles (aerosols) (Zahra et al. 2022). Given the critical nature of this matter, the present study employs Sentinel-5 satellite images to quantify pollution levels in Iran's metropolises. One key advantage of utilizing Sentinel-5 satellite images in pollution monitoring is the ability to obtain comprehensive and global observations of the Earth's surface at a specific moment, facilitating an effective exploration of spatial pollutant distribution. Additionally, this study employs Markov and Cellular Automata (CA)-Markov methods to forecast thermal pollution levels in 2040. Moreover, deep neural networks and linear regression models are employed to predict air pollution levels. The research also identifies factories and industries as the primary sources of air pollution. In summary, the research objectives are as follows:

- Generating maps of O₃, CH₄, NO₂, CO₂, SO₂, CO, and suspended particles (aerosols) by utilizing Sentinel-5 satellite data.
- Analyzing satellite MODIS images to assess Land Surface Temperature (LST) and Normalized Difference Vegetation Index (NDVI) for vegetation evaluation.
- Employing Markov and CA-Markov methods, along with deep neural networks, to forecast future air pollution levels.
- Utilizing artificial neural network (ANN) techniques such as multilayer perceptron (MLP), radial basis function (RBF), and LSTM to predict air pollution levels.

It's worth noting that this article addresses several deficiencies compared to previous studies. These improvements encompass a comprehensive examination of a broad spectrum of pollutants, including O₃, CH₄, NO₂, CO₂, SO₂, CO, and airborne particles. The research delves into pollution levels in both 2001 and 2019, enabling the long-term comparison of changes through the utilization of Sentinel

5 satellite imagery in the investigation of pollution across industrial cities. Furthermore, it thoroughly explores the connection between pollution and land surface temperature (LST). To enhance prediction accuracy and facilitate effective pollution management, this work employs neural network models, specifically MLP, RBF, and LSTM, for air pollution forecasting.

This essay contributes to a comprehensive understanding of air pollution dynamics, enabling better decision-making and proactive measures for pollution control and management. The paper is organized as follows: Sect. 2 presents the data utilized in this study along with the characteristics of the study area. Furthermore, this section elucidates the methods employed to assess air pollution. Section 3 presents the findings of the research. Lastly, Sect. 4 offers the concluding remarks.

Materials and methods

Study area

The study area spans between 24°18' and 38°12' North and between 47°42' and 61°06' East (Fig. 2). Several major industrial cities in Iran, such as Tehran, Isfahan, Shiraz, Ahvaz, Khorasan, Bushehr, Mazandaran, Qom, Semnan, Yazd, Fars, Khorasan, and Bushehr, were chosen for monitoring air pollutants using Sentinel-5 data (Fig. 2). Detailed information on the area, population, and sources of air pollution for the investigated provinces is presented in Table S1. Notably, the Tehran metropolis boasts the highest population, approximately 13.27 million people, encompassing an area of 18,814 km². On the other hand, Semnan city has the lowest population and covers an area of 97,481 km². The total study area spans 673,193 km² and exhibits diverse climatic conditions. The northern part experiences a humid and rainy climate, while the southern region is characterized by a dry, semi-arid, and hot climate.

Data used in this study

In this study, satellite images were used as input data to prepare spectral, thermal and air-polluting gas indices. Further explanations related to satellite images for the preparation of these indicators are given below.

The spectral and thermal indices (NDVI and LST)

The spectral and thermal indices, specifically the NDVI and LST, were generated using images acquired from the MODIS sensor on board the Terra satellite. The Terra satellite was launched on December 22, 2000, and it houses five sensors, including MODIS. MODIS data consists of 36

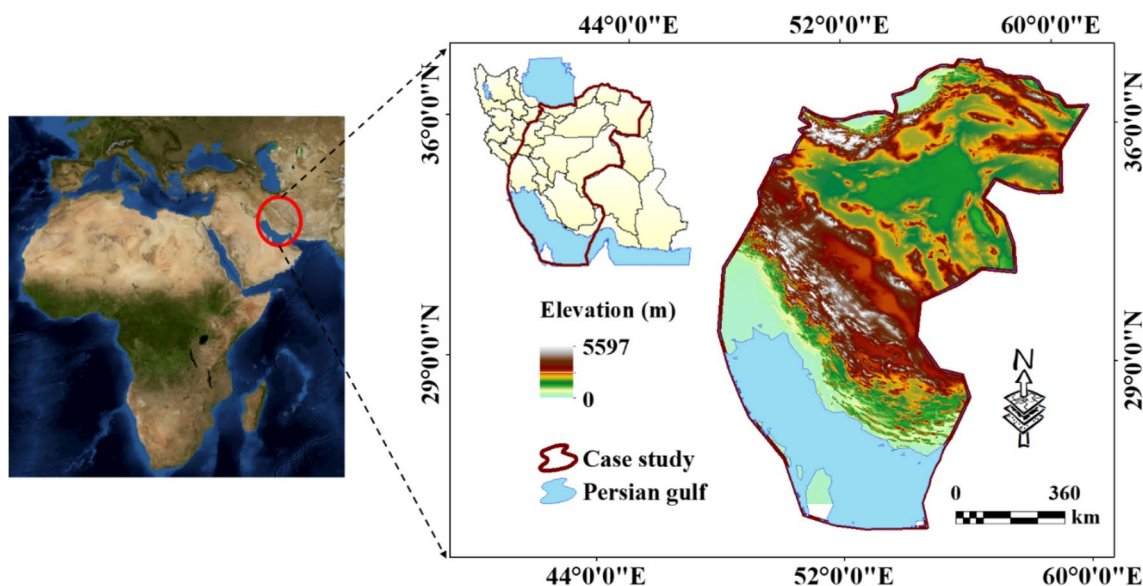


Fig. 2 Position of the study area

spectral bands, encompassing visible, infrared, mid-infrared, short-wave infrared, and thermal wavelengths. While all 36 bands capture data from the Earth's surface, only 20 of them provide information on radiance and reflectance. The satellite's spatial resolution is categorized into three types: Bands 1 and 2 have a resolution of 250 m, bands 3 to 7 have a resolution of 500 m, and the remaining bands have a resolution of 1 km (Campagnolo et al. 2016). The MODIS sensor boasts a 12-bit radiometric resolution and a swath width of 2,330 km, which are among its notable advantages. In this study, bands 1 to 7 of the MODIS sensor were utilized to generate the vegetation index, as outlined in Table S2. By employing these specific bands from the MODIS sensor, researchers were able to derive valuable information regarding vegetation density and health (through NDVI) as well as land surface temperature patterns (via LST). These indices served as essential indicators for assessing environmental conditions and studying the relationship between vegetation, land surface temperature, and air quality in the study area.

Atmospheric pollutants

O₃, CH₄, aerosols, CO, CO₂, NO₂, and SO₂ were detected by using the images of Sentinel-5 satellite. Level three (L3) images of the Sentinel-5 TROPOMI satellite were used to monitor pollutant concentrations. In addition, three emission recording stations in the northwest, two in the center, and two in the west were used to verify the results of air pollution in this study (<https://aqms.doe.ir/>). After preparing the images, various geometric and radiometric corrections were performed on the target image using ENVI software.

Method

Extracting the temperature and vegetation indices using Terra satellite images

To calculate the LST and NDVI indices using the MODIS sensor, the following steps are followed:

Step 1. Converting digital values (DN) to spectral radiation. The MOD11 products from the Terra and Aqua satellites include surface temperature and emissivity coefficient, generated at Level 2 and Level 3, respectively, with a spatial resolution of 1 km and 5 km. For this study, MOD11—Level 2 images from the Terra satellite were utilized. Surface temperature was determined using the separate window algorithm proposed by (Wang et al. 2008), as well as the night and day surface temperature algorithm. Following geometric corrections, the acquired images were transformed into two outputs: surface temperature and emissivity coefficient. Equation 1 demonstrates the conversion of digital values to spectral radiation in the thermal bands of this satellite.

$$L_{\gamma} = \frac{[(L_{max} - L_{min})]}{[(QCal_{max} - QCal_{min}) \times QCal] + L_{min}} \tag{1}$$

where $QCal$ represents the digital value of the image, $QCal_{min}$ represents the minimum gray level value, and $QCal_{max}$ represents the maximum gray level value. Additionally, L_{min} and L_{max} represent the reference spectral radiation values of bands 31 and 32 in DN (Digital Numbers), which are 0 and 255 (Wm⁻²Sr⁻¹ μm⁻¹)

respectively. The specific values of L_{min} and L_{max} can be found in the header file of the images.

Step 2. Converting spectral radiation to spectral reflection. Using Eq. 2, MODIS thermal data were converted from spectral radiation to spectral reflectance by the following formula:

$$TB = \frac{K_2}{I_n \left(\frac{K_1}{L_\lambda} + 1 \right)} \tag{2}$$

where TB is the effective temperature value in the satellite in degrees Kelvin (Wm-2Sr-1 μ m-1), L_λ is the spectral radiance (Wm-2Sr-1 μ m-1), and K_1 and K_2 are constant calibration coefficients in nanometers.

Step 3. Determining surface emissivity. The surface emissivity was estimated using the normalized difference index thresholding method to analyze soil characteristics in each pixel and quantify the emission and the variation in emission. In this approach, pixels with NDVI less than 0.2 were classified as dry soil, with a radiation power value of 0.978. Pixels with NDVI greater than 0.5 were associated with higher vegetation density, and a radiation power value of 0.985 was assigned to them. For pixels with NDVI values between 0.2 and 0.5, which represent a combination of vegetation and soil, the radiation power was determined using Eq. 3.

$$\epsilon = \epsilon_{veg} P_v + \epsilon_{Soil}(1 - P_v) \tag{3}$$

where P_v is the ratio of plant cover which is calculated using Eq. 4.

$$P_v = \left(\frac{NDVI - NDVI_{min}}{NDVI_{max} - NDVI_{min}} \right)^2 \tag{4}$$

where NDVI index is calculated from the ratio of red and near infrared bands (Eq. 5).

$$NDVI = \frac{\varphi_{band2} - \varphi_{band1}}{\varphi_{band2} + \varphi_{band1}} \tag{5}$$

where ρ is the reflection coefficient of multispectral bands (Eq. 6)

$$\rho = \frac{\pi \times L_\lambda \times d^2}{ESUN_\lambda \times \theta_s} \tag{6}$$

where L_λ is the spectral radiance, d is the distance from the earth to the sun, $ESUN_\lambda$ is the average solar reflectance and θ is the sun's zenith angle in degrees.

Step 4. Calculating LST index. To calculate LST was used Signal channel method (Eq. 7).

$$T_s = \gamma [\epsilon^{-1}(L_{sensor} + \psi_2) + \psi_3] + \delta \tag{7}$$

In this context, the emissivity of the Earth's surface (ϵ) is determined using the given relationship. The atmospheric effect regulator (ψ) is applied to the thermal images. The thermal radiance (L_{sensor}) is calculated for band 10. The parameter δ depends on the Planck function, and γ represents the wavelength at which the detector operates. All of these parameters are calculated using Eqs. 8, 9, 10 and 11.

$$\psi_2 = 0.00918w^2 + 1.36072w + 0.27514 \tag{8}$$

$$\psi_3 = 0.00918w^2 + 1.36072w - 0.27515 \tag{9}$$

$$\gamma = T_{sensor}^2 / b_\gamma \times L_{sensor} \tag{10}$$

$$\delta = T_{sensor} - T_{sensor}^2 / b_\gamma \tag{11}$$

where L_{sensor} is the radiance of the thermal band, T_{sensor} is the temperature of the radiance of the thermal band, and both parameters were calculated for band 10. b_γ is a constant number that is equivalent to 1,324 for band 10.

Finally, LST is calculated based on Eq. 12:

$$LST = \frac{T}{\left(+ \left(y \frac{T}{\frac{hxc}{s}} \right) x \log(\epsilon) \right)} \tag{12}$$

Multi linear regression

The linear regression method was employed to examine the correlation between vegetation quantity, surface temperature, and pollutant levels. Linear regression is recognized as one of the most commonly used techniques for data and information modeling. In this study, a simple regression model was utilized with the dependent variable Y (surface temperature, LST) and independent variables (such as SO₂, NO₂, CO₂, CO, etc.) denoted as X₁, X₂, ..., X_{p-1}. The model can be defined as follows (Eq. 13):

$$Y_i = \beta_2 X_1 + \beta_2 X_2 + \dots + \beta_{p-1} X_{p-1} + \epsilon_i \quad i = 1, 2, \dots, n \tag{13}$$

This Eq. 14 can also be expressed as a matrix:

$$Y_{n \times 1} = \begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix} X_{n \times p} = \begin{bmatrix} 1 & X_{11} & X_{12} & \dots & X_{1,p-1} \\ 1 & X_{21} & X_{22} & \dots & X_{2,p-1} \\ \vdots & \vdots & \vdots & \dots & \vdots \\ 1 & X_{n1} & X_{n2} & \dots & X_{n,p-1} \end{bmatrix} \tag{14}$$

where the matrix X shows the observed values of $p-1$ variables for n samples. The Y vector is the observed value of the dependent variable for n samples.

Markov and CA-Markov chains

The Markov chain method was used to predict thermal pollution in the coming years. The Markov chain is based on developing a transition probability matrix for a parameter over time. Based on the Markov chain model, this matrix is used to estimate the probability of a pixel switching from one class to another over time. In the method, when a chain is considered a random process, its value at time t is determined only by its value at time $t-1$, i.e. X_t (Eq. 15). In a given period of time, the relation $P\{X_t = \alpha_j | X_{t-1} = \alpha_i\}$ calculates the probability that a process will transition from state α_i to state α_j . The probability of transition of n steps is shown as P_{ij} when n steps are required to make this transition. First-order homogeneous Markov chains are shown in Eq. 16.

$$P\{X_t = \alpha_j | X_0 = \alpha_0, X_1 = \alpha_1, \dots, X_{t-1} = \alpha_i\} = P\{X_t = \alpha_j | X_{t-1} = \alpha_i\} \tag{15}$$

$$P\{X_t = \alpha_j | X_{t-1} = \alpha_i\} = P_{ij} \tag{16}$$

where P_{ij} is the transition probability matrix

The use of Markov chains alone does not allow for the determination of spatial distribution changes. While the transfer probabilities for each class in the model may be highly accurate, the spatial distribution of the classes remains unknown. To address this issue, this study employed the CA-Markov chain model. This model integrates cellular automata, Markov chains, and multipurpose land allocation (MOLA) to predict future land use changes. In order to assess the accuracy of the Markov and CA-Markov methods' results, the Kappa method was utilized (Eq. 17).

$$Kappa = \frac{p_o - p_c}{1 - p_c} \times 100 \tag{17}$$

where O is correctly observed and PC is expected agreement.

Prediction methods

Multilayer Perceptron (MLP) method: The MLP method is a type of feedforward artificial neural network. It is comprised of three layers: input, hidden, and output. In this network, all nodes, except for the input nodes, are neurons that utilize nonlinear activation functions. The MLP utilizes a supervised learning technique known as feedback for training. It is widely employed for predicting and solving nonlinear problems. The Back Propagation algorithm is commonly utilized to train these networks. The input values for each neuron are calculated using Eq. 18.

$$net_i^n = \sum_{j=1}^m w_{ji}^n O_j^{n-1} \tag{18}$$

where net_i^n is the input values of the i th neuron in the n th layer, w_{ji}^n is the connection weight between the i th neuron in the n th layer and the j th neuron in the layer $n-1$, O_j^{n-1} is the output of j th neuron in layer $n-1$ and M is the number of neurons in layer $n-1$.

The values calculated from Eq. 18 are converted into numerical values in each neuron. A sigmoid function is commonly used for this purpose, as defined by Eq. 19:

$$Sig\left(\frac{net_i^n}{j}\right) = \frac{1}{1 + Exp(-net_i^n)} \tag{19}$$

Using this equation, each neuron's calculated output is multiplied by its weight matrix before moving onto the next step. In the final step, the calculated output of the network is compared with the actual output.

Radial Basis Function (RBF) method The RBF networks are feedforward networks with an intermediate layer. Fig. S1 illustrates a RBF. The training of RBF networks generally consists of two stages. The first part consists mainly of unsupervised learning that clustering, the parameters of the base functions (centers and widths) are determined using the input information. In the second part, which is supervised learning, the weights between the intermediate layer and the output layer are determined using slope reduction and linear regression methods. The middle RBF neuron is connected to each of the input neurons with weight parameters. These parameters are the centers of neurons. The output of each intermediate neuron depends on the distance between the input vector ($X = [x_1, x_2, \dots, x_n]$), and the radial center vector ($W = [w_{1j}, w_{2j}, \dots, w_{nj}]$), which is defined as Eq. 20.

$$\delta = \sqrt{\sum_{i=1}^n (x_i - w_{ij})^2} \tag{20}$$

The output of the intermediate neuron was determined using the Gaussian function, which is calculated as Eq. 21:

$$f(\delta_j) = Exp(-\lambda \delta_j^2) \tag{21}$$

where λ is a constant coefficient.

Finally, the outputs of the output layer are calculated from Eq. 22:

$$Z_k = \sum_{j=1}^j b_{jk} y_j \tag{22}$$

where b_{jk} is the weight coefficient between the j th neuron of the middle layer and the k th neuron of the output layer and y_j is the output of the j th neuron of the middle layer.

Long-term short-term memory (LSTM) method The LSTM is a special type of Recurrent Neural Network (RNN) that is capable of learning long-term dependencies. Fig. S2 illustrates four interactive layers in an LSTM cell (Pal et al. 2018).

As shown in Fig. 3, x is the input, h and C are memory vectors. C are cell activation vectors, all of which equal the hidden vector h . σ is a logistic sigmoid function. \tanh puts values between -1 and 1. As a first step, the "forget gate layer" determines which information to remove. Second step, the "input gateway layer" determines which values need to

be updated. Then, a \tanh layer creates a vector of new candidate values. In the update operation, \overline{C}_t updates the state of the old cell C_{t-1} to the state of the new cell C_t . After that, the output gate layer decides which parts of the cell state to output. Equations 1, 2, 3, 4, 5, 6, 7 and 8 represent the equations calculated by a cell (Eqs. 23, 24, 25, 26 and 27).

$$\text{Forgetgate layer}(f_t) = \sigma(w_f \cdot [h_{t-1}, x_t] + b_f) \tag{23}$$

$$\text{Inputgate layer}(i_t) = \sigma(w_i \cdot [h_{t-1}, x_t] + b_i) \tag{24}$$

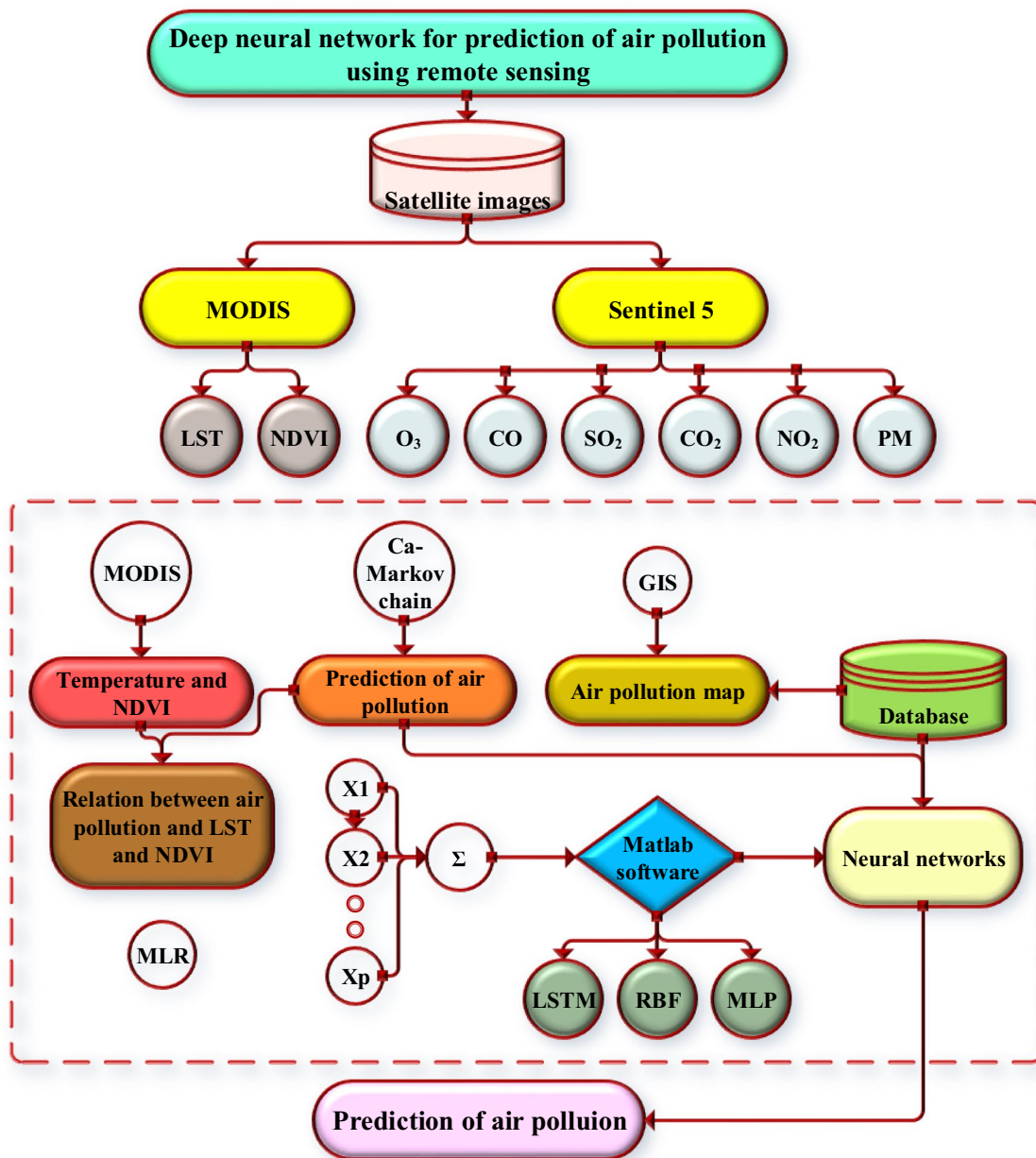


Fig. 3 A graphical representation of research steps

$$\text{Newmemorycell}(\tilde{C}_t) = \tanh(w_f \cdot [h_{t-1}, x_t] + b_f) \quad (25)$$

$$\text{Finalmemorycell}(C_t) = f_t \times C_{t-1} + i_t \times \tilde{C}_t \quad (26)$$

$$\text{Outputgatelayer}(o_t) = \sigma(w_o \cdot [o_t + \tanh(C_t), x_t] + b_o) \quad (27)$$

where w is a matrix, b is the bias, i is the input gate, f is the forget gate, and o is the output gate.

Validation of models

Mean Absolute Percentage Error (MAPE) was employed to evaluate the accuracy of pollution maps generated from satellite images for each pollutant, while Root Mean Square Error (RMSE) was utilized to assess the accuracy of the neural network models (Jiang et al. 2022a, b; Wang et al. 2022) (Eqs. 28 and 29).

$$RMape = (1/n) \times \sum |(A - P)/A| \times 100\% \quad (28)$$

where A is the actual or observed vegetation index values from ground measurements or reference data. P is the vegetation index values predicted or estimated using remote sensing data or models. N is the number of data points (samples) in the dataset.

$$RMSE = \sqrt{\frac{\sum_{t=1}^n (A_t - F_t)^2}{n}} \quad (29)$$

where A_t is the actual value of temperature in meteorological stations, F_t is the predicted value obtained from separate window algorithms, and N is the total number of meteorological stations (Guo and He 2021).

The research steps to investigate air pollution are shown in Fig. 3.

It is worth noting that in the selection of MLP, RBF, and LSTM as neural network models for pollution prediction, we carefully considered their unique advantages within the study context. MLP, for instance, offers a high degree of versatility, enabling it to effectively capture intricate relationships within the data, rendering it well-suited for our multifaceted pollutant predictions. Additionally, RBF excels in modeling non-linear relationships, a pivotal feature when dealing with the intricacies of environmental data. Furthermore, LSTM's proficiency in handling time-dependent patterns significantly enhances our capacity to predict pollution levels over temporal scales. Amidst the diverse array of neural network architectures available, our decision to employ MLP, RBF, and LSTM stems from their alignment with our research objectives,

ultimately affording us the ability to attain superior accuracy and robust predictions in the domain of air pollution monitoring.

Results

Spatio-temporal evaluation of spectral indices using MODIS sensor

The NDVI index was employed to evaluate the vegetation status in the study area, while the LST index was used to assess the land surface temperature. The analysis was conducted for the years 2001, 2006, 2010, 2015, and 2019 during the spring, summer, autumn, and winter seasons. Fig. S3 illustrates the minimum, maximum, and average values for each examined year and month. Based on the findings presented in Fig. S3, it is evident that the vegetation amount decreased in 2019 compared to 2001, while the temperature exhibited an increase. Specifically, during the spring season, the LST index reached its highest value of 59.91 in 2019, and its lowest value of -16.95 in 2006. Furthermore, Fig. S3 demonstrates that industrial cities like Yazd, Isfahan, Bush-ehr, the southern region of Fars province, Khuzestan, and Tehran have recorded the highest LST values. This increase in LST values can be attributed to the expansion of industrial activities, which have had a significant impact on the temperature (Wang and Wei 2020; Zhang et al. 2022).

The spatial and temporal distribution of these indices can be observed in Fig. S4 and Fig. S5. Fig. S4 illustrates that the highest values of the indices are concentrated in the northwestern and eastern halves of the region, indicated by green and yellowish-green colors with values exceeding 0.8. Conversely, the lowest values of the index, represented by red areas, are predominantly located in the eastern and central regions where vegetation is scarce or absent. The index values exhibit a seasonal variation, with lower values observed during autumn and winter compared to spring and summer.

The NDVI index values for 2019 have decreased compared to 2010 and 2001, reaching values below 0.2, which indicates a reduction in vegetation. Industrial activities in these areas have contributed to the decline in the index values and further degradation of vegetation (Zeng and He 2019). On the other hand, the LST index reflects the amount of heat on the Earth's surface, serving as an indicator of thermal pollution. The lowest LST values are observed in the northwestern and central regions, where vegetation cover is most prominent. Conversely, the central and southern regions exhibit the highest LST values, indicating elevated temperatures.

Atmospheric pollutants

In this study, Sentinel 5 images were utilized to monitor air pollutants in 17 provinces over a span of two years, encompassing the periods before and during the COVID-19 pandemic. The average levels of various air-polluting gases, such as CH₄, CO, CO₂, NO₂, O₃, SO₂, and aerosols, were measured during the months of February, November, August, and May from 2018 to 2021. Fig. S6 presents the mean, maximum, minimum, and standard deviation values for each pollutant. The results indicate that the highest concentrations of these gases were observed in August, while the lowest levels were recorded in February. Additionally, the study found that these pollutants were less abundant during the COVID-19 pandemic, attributed to a decrease in industrial activities.

The estimated maximum amounts of O₃, SO₂, CO, and CO₂ were 0.21 ppm (November), 0.0088 ppm (February), 0.062 ppm (February), and 0.00083 ppm (November) respectively, while the minimum levels were 0.105 ppm, 0 ppm, 0 ppm, and 0 ppm respectively. For aerosols, CH₄, and NO₂, the maximum values in May and February were 3.94 (August), 2004.91 nanomol/mol, and 0.0022 ppm respectively, while the minimum levels were 0, 1757 nanomol/mol (February), and 0 respectively. Figs. S7 to S13 present the spatio-temporal analysis of each of these gases. Based on Fig. S7, it is evident that the maximum concentration of CH₄ gas in August 2021 reached 2004.98 nanomol/mol in the northern (Tehran) and eastern (Yazd) parts of the study area. In February 2019, the western regions of the study area recorded the minimum concentration of this pollutant at 1757.19 nanomol/mol. There are significant areas in the south and small portions in the west where this pollutant is absent. The changes in CH₄ levels from 2019 to 2021 indicate a decreasing trend in most areas and months. Notably, the concentration of CH₄ has also decreased in November and February 2021, likely due to the quarantine measures and reduced industrial activities during the COVID-19 outbreak. There is a clear correlation between industrial activities and the presence of CH₄ in the atmosphere. CH₄ is a greenhouse gas that contributes to global warming and can persist in the atmosphere for an extended period (Wang et al. 2021a).

In February 2019, the highest concentration of CO was recorded at 0.051 ppm in small areas located in the northern and southern parts of the study area, as shown in Fig. S8. The minimum concentration of this pollutant in August 2021 was 0.005 ppm in the northern, central, and western regions. The analysis reveals a decreasing trend in the levels of CO over time, leading to a decrease in its atmospheric distribution in August 2021.

CO is produced through the incomplete combustion of carbon. It is a highly toxic gas that lacks specific color or

odor. CO poses a significant threat to human health and can be lethal. When inhaled, it combines with hemoglobin in the blood to form carboxyhemoglobin, which hinders oxygenation in the body and can ultimately lead to death. Human activities, such as metal smelting, industrial processing plants, chemical industries, the use of fossil fuels, and even vehicles, contribute to the release of CO into the atmosphere (Prockop and Chichkova 2007; Oves et al. 2018).

Fig. S9 illustrates the spatio-temporal distribution of CO₂ in the studied area. The highest concentrations of this gas were observed in the northern regions (Golestan and Tehran) and the southern regions (Khuzestan and Bushehr) in August 2019. In February 2021, parts of the northern and western areas of the region exhibited no pollution. The changing trend of CO₂ indicates a decreasing pattern in the region, with the maximum concentrations observed in highly polluted cities like Tehran or Khuzestan, which are major consumers of fossil fuels.

CO₂ is primarily released into the atmosphere through the combustion of fossil fuels such as coal, oil, and gas, as well as the burning of organic materials. The accumulation of CO₂ in the atmosphere contributes to the depletion of the protective ozone layer and leads to an increase in the Earth's temperature. This phenomenon, known as global warming, is responsible for the melting of glaciers and the escalation of flood events (Soeder 2021; Roy et al. 2022).

Fig. S10 depicts the spatio-temporal distribution of NO₂. The highest concentrations of this gas were observed in Tehran (northern region), Ahvaz (western region), Isfahan and Yazd (central region), and Bushehr (southern region). The highest recorded level of NO₂ was in Tehran in May 2019, reaching 0.0021 ppm. On the other hand, the southern and southeastern regions exhibited the lowest concentrations of this pollutant. The decrease in NO₂ levels can be attributed to the quarantine measures implemented during the COVID-19 pandemic, which resulted in a reduction in industrial activities. NO₂ is a significant air pollutant primarily emitted from the consumption of fossil fuels. Its reaction with H₂O forms nitric acid, which can cause severe corrosion of metals (Li et al. 2019a). At high concentrations, NO₂ contributes to the formation of intense smog and reduces visibility. Individuals with asthma may experience worsened respiratory symptoms due to the presence of this pollutant.

Fig. S11 illustrates the spatio-temporal distribution of O₃ in the study area. The results indicate a decrease in the concentration of this gas in 2021 compared to 2019, with a decreasing trend over time. In February 2019, the highest level of O₃ was recorded in Tehran, reaching 0.163 ppm. Conversely, the lowest level of O₃, 0.103 ppm, was observed in the northern areas (Golestan province) in November 2020.

O₃ is a highly reactive and strong oxidizing gas, making it a hazardous pollutant. The maximum acceptable standard for O₃ concentration is 0.125 ppm in air volume, and values

exceeding this threshold are considered very dangerous, particularly on hot and sunny days. Human health, especially that of the elderly, children, and individuals with respiratory conditions, is significantly affected by this gas (Luong et al. 2018). Various pollution sources contribute to the release of O_3 into the atmosphere, including vehicle emissions, power plants, refineries, and petrochemical plants. Tehran, with its higher industrial activities and traffic volume, has a higher likelihood of experiencing air pollution from O_3 compared to other areas (Tondelli et al. 2022).

Fig. S12 presents the spatio-temporal values of SO_2 gas during the seasons of November, February, August, and May from 2018 to 2021. The results indicate that this gas is most prevalent in the southern region, particularly in Khuzestan province, and least prevalent in parts of the north. In February 2019, Tehran, Khuzestan, and Bushehr provinces recorded the highest concentrations of SO_2 . Conversely, in February 2021, the lowest levels of this gas were observed in the northwestern parts of the region. Over the studied time period, there has been a decrease in the concentration of this pollutant, which can be attributed to the implementation of quarantine measures during the COVID-19 pandemic. SO_2 typically remains in the atmosphere for approximately two to three days. Research indicates that 90% of sulfur oxides, including SO_2 , are produced by the combustion of fossil fuels (Park et al. 2021). Cities like Khuzestan, with higher levels of fossil fuel consumption, tend to exhibit higher concentrations of this gas compared to other regions. Patients with respiratory diseases are particularly vulnerable to the adverse effects of SO_2 exposure.

Fig. S13 illustrates the spatial-temporal distribution of aerosol concentrations. The results indicate that the minimum amount of aerosols is observed in the northern regions, specifically in Golestan province and the Persian Gulf waters, while the maximum amount is found in the southern and southeastern parts of the study area. In August 2019, the highest concentration of aerosols reached 3.194 ppm.

To assess the accuracy of each pollution map generated from satellite images, MAPE values were calculated at 10 specific locations within the region where ground station measurements of pollution values were taken. The results indicated MAPE values of 3% for CH_4 , 5% for CO, 2% for CO_2 , 4% for NO_2 , 8% for O_3 , 3% for SO_2 , and 5% for aerosols, respectively. Based on the MAPE values, which fall within the "very good" classification, it can be concluded that the accuracy of the satellite images is sufficiently high for determining the pollution levels.

The trend of aerosol concentration shows a reduction over time, which can be attributed to the decrease in human activities during the COVID-19 pandemic. Increased concentrations of suspended particles in the atmosphere contribute to higher aerosol levels. These particles, which can include substances such as silica, asbestos, or diesel, pose

risks to human health. Inhalation of these particles can lead to various respiratory problems. Research indicates that a significant portion of aerosols is composed of sulfate particles originating from fossil fuel combustion (Wang et al. 2021b). Aerosols also have climate implications, as they can block sunlight from reaching the Earth's surface, leading to cooling effects. Furthermore, when these particles enter the respiratory system, they have the potential to cause cancer (Nho 2020).

Figure 4 provides information on the concentrations of various pollutants in different cities. According to Fig. 4, Tehran had the highest methane (CH_4) concentration at 2,004 nanomol/mol, while Ahvaz had the lowest concentration at 1,900 nanomol/mol. Tehran, Ahvaz, Asaluyeh, and Bushehr recorded the highest levels of carbon dioxide (CO_2), nitrogen dioxide (NO_2), ozone (O_3), and sulfur dioxide (SO_2), respectively. Ahvaz had the highest aerosol concentration at 3.94 ppm. Khuzestan province, particularly the Ramin region, exhibited the highest CO_2 levels, which can be attributed to the presence of multiple power plants in the area. The Ramin region is home to one of the largest thermal power plants globally, with a capacity of 1,903 MW. Khuzestan relies on this power plant for 42% of its electricity needs, while the entire country relies on it for 6%. The Fig. 4 also indicates that cities like Bandar Abbas, Tehran, and Ahvaz had high carbon monoxide (CO) levels. These cities experience increased levels of CO due to car fuel combustion and emissions from smoke. In Ahvaz, the majority of suspended particles in the atmosphere are caused by severe wind erosion and the subsequent increase in suspended particles.

Figure 5 presents the results of the total concentration of seven pollutants, which were spatially normalized and combined to create a final air pollution map. According to the map, Tehran and Bushehr were identified as the most polluted cities during the examined period. The overall trend indicated a decrease in pollution levels in 2021 compared to 2019. This reduction can be attributed to the impact of quarantine measures implemented during the Coronavirus period, resulting in reduced industrial activities and fewer vehicle trips. To create the final air pollution map, each pollution map was normalized within the GIS environment, ranging between 0 and 1. These normalized maps were then summed to generate the final air pollution map. The results highlight that the city, particularly industrial areas, exhibited higher levels of pollution compared to other regions. This observation reinforces the strong connection between air pollution and industrial activities.

Fig. S14 illustrates the relationship between air pollution and high traffic congestion, indicating that areas with high pollution levels often coincide with areas of heavy traffic. The presence of a large number of cars on the roads can contribute to increased emissions of pollutants such as

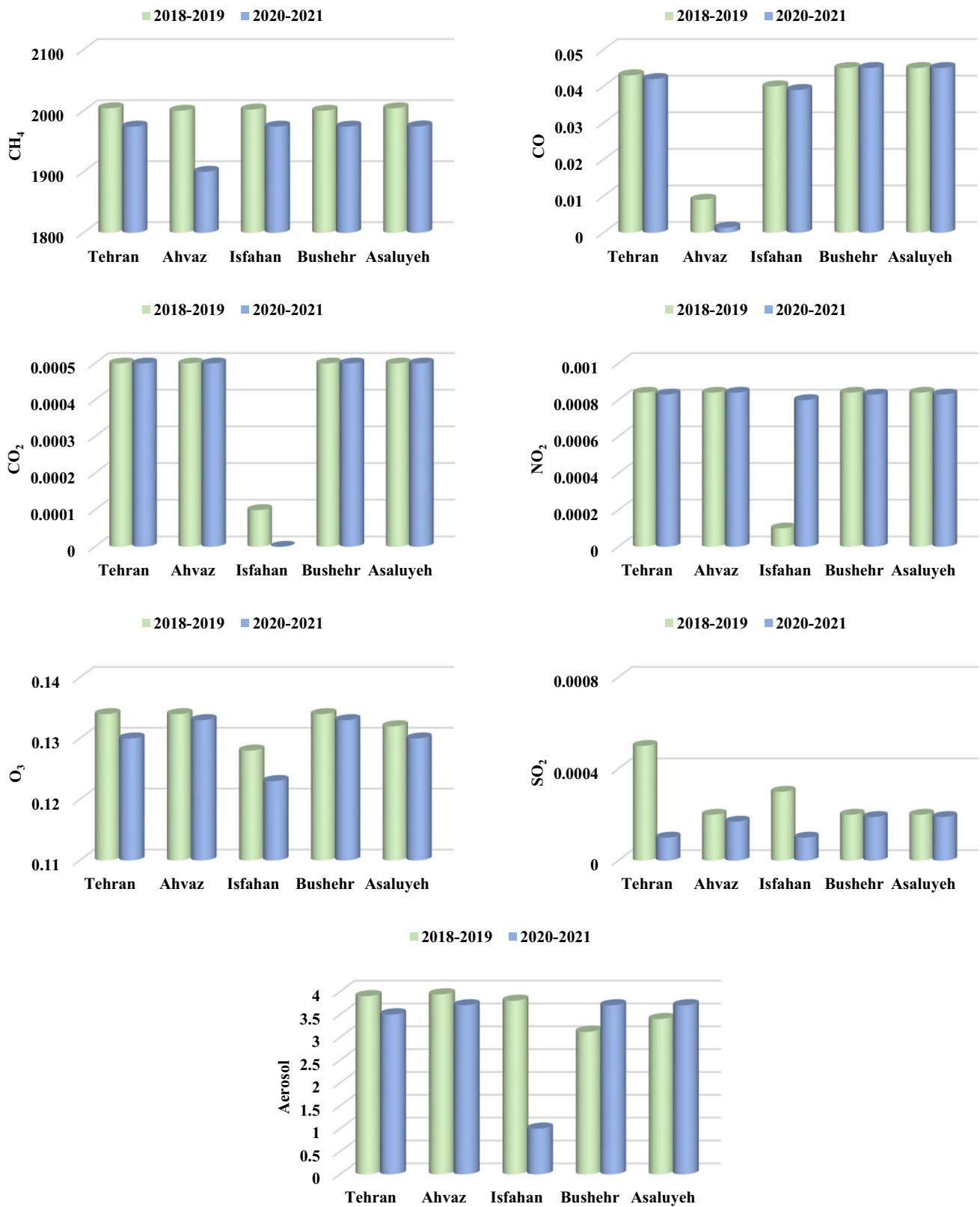
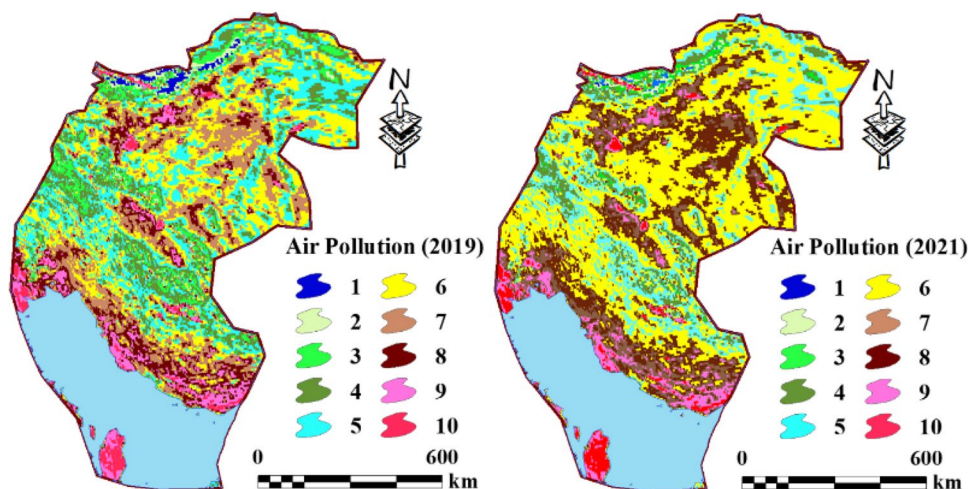


Fig. 4 Comparison of air pollution level of 17 studied cities in 2018–2019 and 2019–2020

Fig. 5 Final maps of air pollution in 2019 and 2021



carbon monoxide (CO), nitrogen dioxide (NO₂), and particulate matter (PM), which are known to be harmful to human health and the environment.

Three emission recording stations in the northwest, two in the center, and two in the west were used to verify the results of air pollution in this study (<http://aqms.doe.ir/home/AQI>). Sentinel 5 satellite images were found to be highly correlated with ground station values ($R^2_{SO_2}=0.81$, $R^2_{CO}=0.81$, $R^2_{O_3}=0.95$, $R^2_{NO_2}=0.9$). Thus, Sentinel-5 satellite images can be used for pollution monitoring and air quality management. Also, the results showed that the most pollution is in the areas with the most industrial activities. There is an accumulation of residential lands and industrial centers in parts of the northwest (Tehran province), east (Khuzestan province), west (Isfahan province), and south (Bushehr and Asaluyeh provinces), as shown in Fig. 6. There is a lot of air pollution in these areas, which can be attributed to the increase in industrial activities.

Relationship between thermal index and pollutants

Fig. S15 presents the results of the correlation analysis between land surface temperature (LST) and air pollution levels in 100 locations. The correlation coefficients (R² values) indicate the strength and direction of the relationship between LST and three pollutants: O₃, CO₂, and CO. According to the results, O₃ has the highest correlation with LST, with an R² value of 0.969. This indicates a strong positive relationship between O₃ levels and land surface temperature. Similarly, CO₂ and CO show moderate positive correlations with LST, with R² values of 0.72 and 0.728, respectively. These findings suggest that as land surface temperature increases, the levels of O₃, CO₂, and CO tend to rise as well. The positive correlation implies that higher temperatures can contribute to the formation or accumulation of

these pollutants in the atmosphere. Table 1 likely provides additional statistical values related to the correlation analysis between LST and air pollution. The R² value of 0.942 indicates a strong positive relationship, and the F statistic of 0.0001 suggests that the relationship is statistically significant. These values support the conclusion that there is a significant and positive association between LST and air pollution levels.

Predicting LST using ANNs

Predicting LST using Markov and CA-Markov

In this section, the study utilizes Landsat satellite data from MODIS to analyze land surface temperature (LST) maps in 2001 and 2019 and predict future LST values for 2040. The Cellular Automata (CA) model is employed with 20 iterations to predict changes in the LST index for the future. The probability matrix in Table S3 represents the distribution of LST classes for each season in 2001 and 2019, while Fig. 7 displays the LST index map for 2040. Based on the observed changes and the projected increase in LST for future years, several indices are expected to be affected. Table S3 highlights the most significant changes in LST values from 2001 to 2019 across different seasons. For instance, the transition from class 1 to 2 shows a 47% increase in LST values during spring. In autumn, transitions from class 1 to 2 and from class 4 to 5 account for 25% of the changes. In summer, a 41% increase is observed in the transition from class 4 to 5, while in winter, the transition from class 1 to 2 represents a 53% increase in LST values. These changes indicate a consistent trend of increasing temperatures across different seasons. The study suggests that the CA-Markov method can be utilized to predict future changes in land cover and guide effective land management strategies.

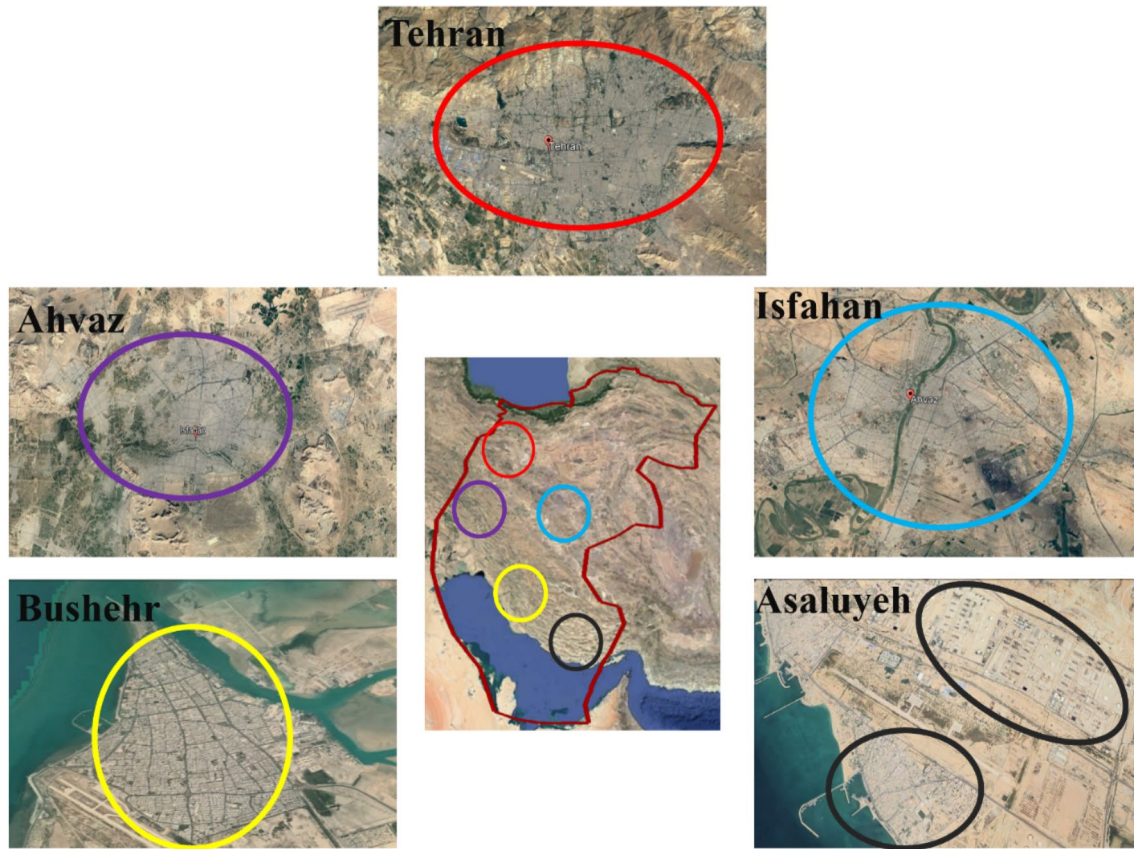


Fig. 6 Land use status in polluted areas

Table 1 Accuracy of Pearson correlation model

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Change Statistics				
					R Square Change	F Change	df1	df2	Sig. F Change
1	.971 ^a	.942	.936	1.072808154026374	.942	142.038	7	61	.0001

a. Predictors: (Constant), Aerosol, CH₄, NO₂, CO₂, SO₂, CO, O₃

Figure 8 provides a visualization of the projected LST values for 2040, indicating the anticipated temperature increase in urban areas. Consequently, it emphasizes the need for appropriate management practices in these areas to mitigate the removal of vegetation, the expansion of urban land use, and the proliferation of bare land. These measures aim to prevent adverse environmental impacts associated with urbanization and promote sustainable land development.

Predicting LST using MLP and RBF method

In this study, the relationship between LST and CO and O₃ indices is examined. The prediction of these parameters is performed using MLP and RBF neural networks (Table 2).

The accuracy of each method is evaluated and presented in Tables 5, along with network features such as R² (coefficient of determination) and RMSE. Additionally, Fig. 8 provides a visual representation of the accuracy comparison. Results indicate that the MLP network outperforms the RBF network in the investigated seasons when comparing the two methods. The MLP method, with 3 neurons in the first hidden layer, and the RBF method, with 4 neurons in the hidden layer, demonstrate the highest accuracy. It is observed that the number of neurons in the hidden layer plays a crucial role in determining the accuracy of both MLP and RBF methods. In this regard, the RBF neural network with 4 neurons in the first hidden layer yields the most accurate structure for LST prediction, exhibiting an R² value of 0.99 and an RMSE value of 1. While the RBF method shows higher

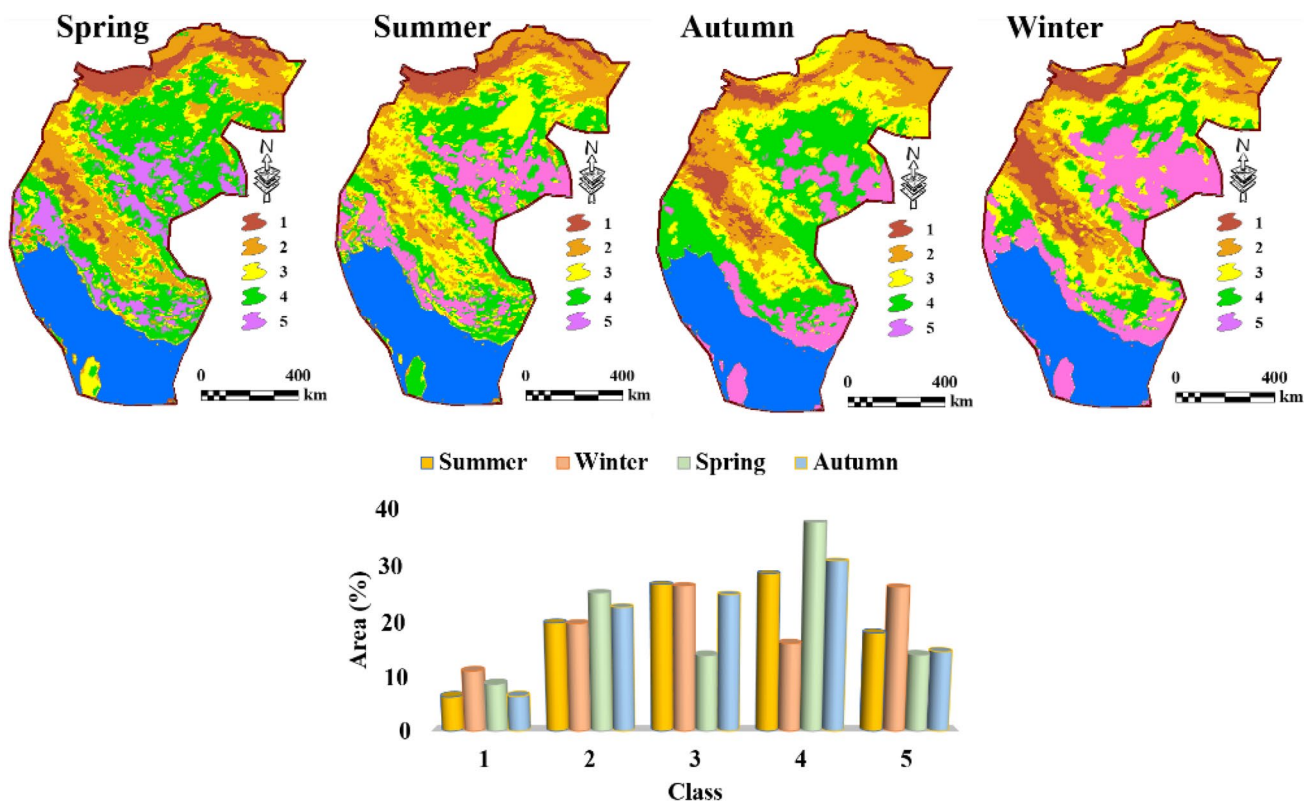


Fig. 7 Area of LST indices predicted using Markov and CA-Markov chains for four seasons in 2040

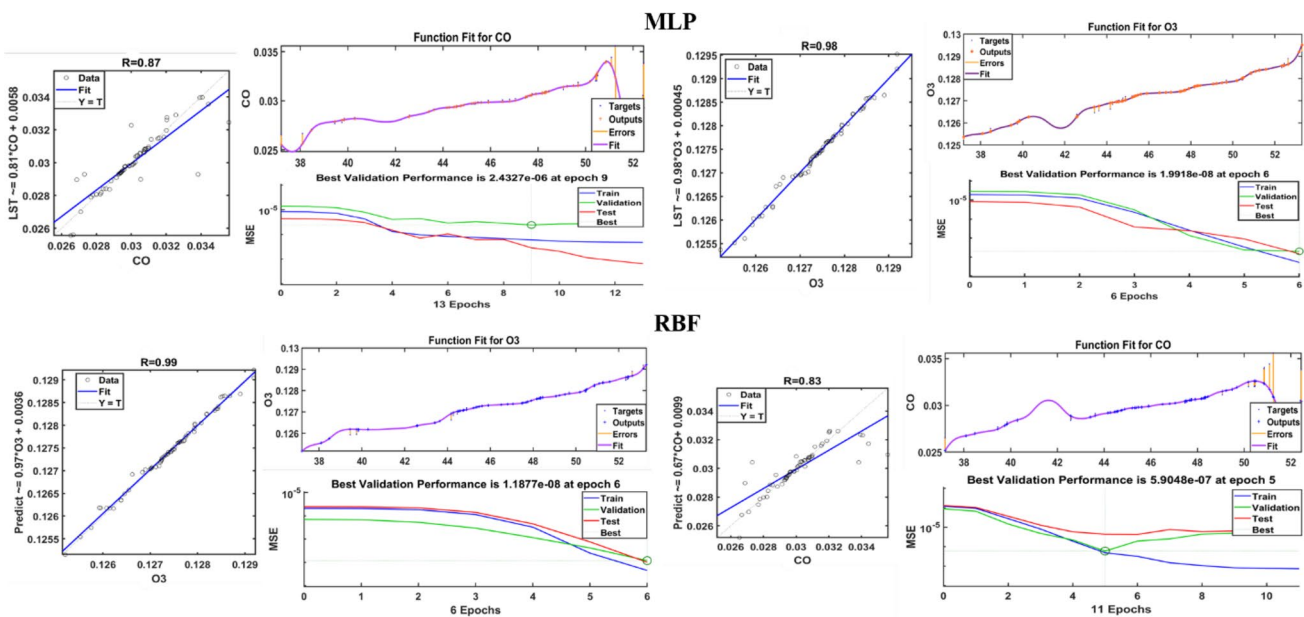


Fig. 8 Performance of the model (a): MLP, (b): RBF

R^2 and lower RMSE, suggesting greater consistency with real-world observations, the MLP method also demonstrates good accuracy in predicting LST values. To further enhance

the accuracy of the network, increasing the number of hidden layers to 3 or 4 may be considered (Souza et al. 2022).

Table 2 Accuracy of ANN models

Models	Input	Model structures	Training		Validation	
			RMSE	R ²	RMSE	R ²
MLP	CO	1:4:1	1	0.98	1	0.99
	O ₃	1:3:1	2	0.95	1	0.98
RBF	CO	1:3:1	3	0.88	3	0.87
	O ₃	1:2:1	3	0.86	4	0.83

Predicting LST using LSTM method

In this section, LSTM neural networks were employed to predict temperature values while considering the influence of industrial activities and environmental pollution. LSTM networks, which are a type of recurrent neural network, were utilized to analyze and forecast time-series data (Vidal and Kristjanpoller 2020; Yadav et al. 2020). Unlike MLP and RBF methods, LSTM networks utilize memory blocks instead of neurons, allowing for better performance (Chang et al. 2018). The LSTM architecture includes different input and output gates to control the state of these memory blocks. In this study, the *tanh* function was used in LSTM to assess the impact of data scaling on prediction. Prior to the training and testing stages, the data were normalized to a range between 0 and 1 and then fed into the network for predicting the last value. Similar to the RBF and MLP models, 30% of the data was allocated for testing and 70% for training. The offline stage of the LSTM network was configured accordingly, while the online settings determined the size of categories and stages for training and prediction (Laib et al. 2019). To evaluate the accuracy of the LSTM model, R² (coefficient of determination) and RMSE were employed. In this particular implementation, the LSTM network consisted of 6 neurons and a cluster size of 10. The input layer, representing LST values, consisted of one neuron (one memory block). LSTM blocks were selected with sigmoid activation functions, and the batch size was set to 1. The number of training epochs was limited to 120, as longer training times did not significantly improve accuracy. The LSTM model was then used for forecasting, and its accuracy was assessed

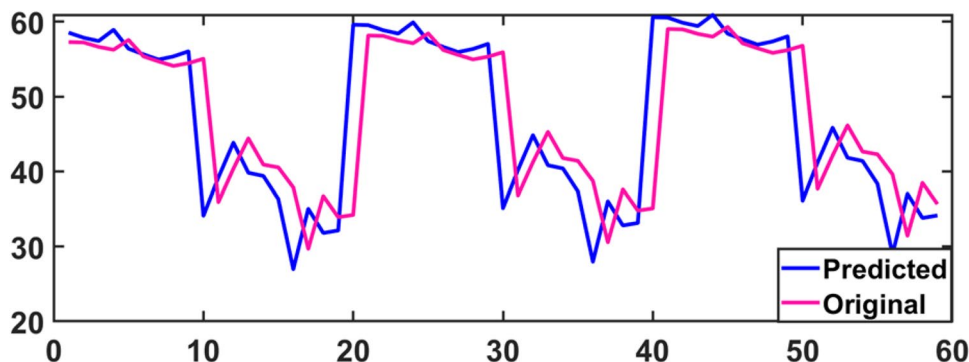
by comparing the predicted data with the original dataset. Figure 9 visualizes the original dataset in pink and the predicted data in blue, demonstrating the comparison between the two. Additionally, Table 3 presents the accuracy results of the LSTM method for testing and training phases. The outcomes of Table 3 indicate the high accuracy of the LSTM model in predicting temperature values in the studied area.

Discussion

In comparing the results of this study with findings from prior research, it becomes evident that the utilization of the NDVI and LST indices has yielded critical insights into the spatial and temporal dynamics of vegetation and land surface temperature in the study area. To put these insights into context, it is essential to examine the contrasts between the present study's observations and those from previous works. The temporal analysis of the NDVI index reveals a concerning trend of decreasing vegetation in the study area over the years, particularly between 2001 and 2019. This decline in NDVI values, signifying reduced vegetation cover, is a cause for alarm and aligns with the findings of Kabiraj (2021), who observed a similar pattern and attributed it to the adverse

Table 3 Accuracy of LSTM method

Input	Training		Validation	
	R ²	RMSE	R ²	RMSE
LST	0.94	6.5	0.98	5

Fig. 9 Results of LSTM method

impacts of industrial activities. This shared observation underlines the long-term ecological consequences of these activities on the region. Conversely, the temporal analysis of the LST index suggests an alarming increase in land surface temperatures, with the highest values concentrated in industrial cities such as Yazd, Isfahan, Bushehr, and others, similar to the findings by Hojamberdiev et al. (2019). These studies have noted a direct correlation between industrial expansion and rising temperatures, substantiating the concerns raised in this research. The implications are far-reaching, as elevated LST values signify the presence of thermal pollution, which can have profound environmental and public health implications.

In comparison to prior research, our study contributes by providing comprehensive and detailed insights into the correlations between LST and multiple key pollutants, which include O_3 , CO_2 , and CO . While existing literature has examined the relationship between LST and individual pollutants, our research delves into the broader context, shedding light on the complex interplay between land surface temperature and a range of pollutants. This enhanced understanding can inform policy and decision-making processes, particularly in the realm of air quality management and climate change mitigation. Our findings corroborate the general trend observed in prior studies that suggest higher temperatures can exacerbate air pollution levels, specifically with regard to O_3 , CO_2 , and CO (Lee et al. 2021). However, the strengths of these correlations and the nuances of these relationships are better elucidated in our study. The robust statistical values and the large-scale, multi-pollutant analysis further contribute to the body of knowledge in the field, making our research a valuable addition to the existing literature.

In addition, this study leveraged advanced neural network models, including MLP and RBF, to analyze the intricate relationship between Land Surface Temperature (LST) and the concentrations of carbon monoxide (CO) and ozone (O_3) indices. Furthermore, it employed LSTM, a recurrent neural network, for predicting temperature values while accounting for the influence of industrial activities and environmental pollution. These methods represent significant contributions to the field of air pollution monitoring, and their results can be compared to findings from prior research for valuable insights (Nanda et al. 2021). The MLP method also displayed commendable accuracy in predicting LST values, indicating the promise of this approach. To further enhance network accuracy, increasing the number of hidden layers to 3 or 4 may be considered, as suggested by previous research (Mohammadi et al. 2020). In contrast, LSTM demonstrated superior performance in modeling temperature values over time, benefiting from its recurrent architecture. These findings collectively contribute to our understanding of the complex dynamics between pollution, temperature,

and prediction methodologies, offering valuable insights for air quality management and decision-making processes. The results of this study showed that neural network methods such as LSTM can accurately predict spectral indices such as the LST index (Shen et al. 2022). The advantage of this method is the combination of stored neurons, which enhances the network's performance. This study found that there is a close relationship between spectral indices and air pollution levels. So, the ambient temperature can be used to predict pollution caused by polluting gases such as SO_2 , O_3 , and CO_2 , which is one of the innovations of this research. Additionally, this study found that thermal pollution is higher in urban areas than in rural areas, which is consistent with the studies of Shi et al. (2021). According to studies (Bera et al. 2020), the reason for the increase in temperature in city areas can be attributed to the type of materials used, traffic, and industrial activities (Baldasano 2020; Müller et al. 2020). Compared to urban areas, there is a lower increase in surface temperature in rural areas due to more vegetation (Duncan et al. 2019; Rahaman et al. 2022; Yao et al. 2019). According to the results of this study, metropolises such as Tehran, Isfahan, and Ahvaz suffer from more air pollution due to overcrowding and industrial activities, which is consistent with studies of Kazemi et al. (2022). This study also shows that air pollution decreased during COVID-19 because traffic and human activities were reduced, which is in line with the results (Barua and Nath 2021; Othman and Latif 2021). Therefore, to reduce pollution in metropolises, managers and politicians should monitor cities extensively to reduce pollution caused by urbanization and human activities.

Conclusion

The findings and contributions of this study hold significant policy implications for addressing the critical issue of air pollution and its impact on climate change. In today's world, where air pollution is a growing concern, the accurate measurement of air pollutants with high spatial and temporal resolution is crucial. Our research, which harnessed remote sensing technologies, satellite imagery, and advanced prediction models, offers valuable insights for policymakers and environmental stakeholders. The identification of Tehran as one of the most polluted cities in both 2019 and 2021 underscores the urgent need for targeted air quality improvement measures in metropolitan areas. The observation of a slight decrease in pollutant levels in 2021, attributed to the Coronavirus pandemic's effects on human and industrial activities, demonstrates the potential for immediate air quality improvements in response to policy interventions during exceptional circumstances. Moreover, our study revealed a decline in vegetation cover and an

increase in LST in 2019 compared to 2001, highlighting shifts in vegetation and temperature patterns over time. The significant correlations found between LST values and specific pollutants, such as O₃, CO, and CO₂, suggest the complex interplay between pollution and local temperature variations. These relationships emphasize the need for comprehensive urban planning and sustainable practices to address these challenges. The predictions made using Markov and CA-Markov chain models, indicating a temperature rise in 2040 followed by increased pollutant levels, serve as a warning of the potential future scenarios if corrective actions are not taken. This information can guide long-term climate and environmental policy planning, helping policymakers prepare for the challenges posed by climate change and air pollution. Lastly, the applicability of advanced neural network models, including LSTM, MLP, and RBF, in accurately predicting pollution levels showcases the potential for incorporating cutting-edge technology in pollution forecasting. This technology-driven approach can enhance the accuracy of early warning systems and aid policymakers in making informed decisions to improve air quality. In conclusion, our research underscores the crucial role of remote sensing technologies and advanced prediction models in monitoring air pollution, temperature, and vegetation dynamics. By leveraging the knowledge gained from these findings, policymakers can develop and implement effective strategies to mitigate air pollution and its associated environmental impacts, leading to a healthier and more sustainable future.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s11356-023-30859-0>.

Acknowledgements It would not have been possible to conduct this research without the support of Shiraz University.

Author's contributions **M.M.:** Conceptualization, Formal analysis, Methodology, Investigation, Project administration, Supervision, Validation, Visualization, Software, Writing—Original draft, Writing—Review & Editing. **F.T.:** Formal analysis, Methodology, Investigation. **T.M.P.:** Supervision, Validation, Writing—Original draft.

Funding The author(s) received no specific funding for this work.

Data availability The datasets used and analysed during the current study are available from the corresponding author on reasonable request.

Declarations

Ethics approval and consent to participate Not applicable.

Consent for publication Not applicable.

Competing interests The authors declare that they have no competing interests.

References

- Al-Ghussain L (2019) Global warming: review on driving forces and mitigation. *Environ Prog Sustainable Energy* 38(1):13–21. <https://doi.org/10.1002/ep.13041>
- Alimissis A, Philippopoulos K, Tzanis CG, Deligiorgi D (2018) Spatial estimation of urban air pollution with the use of artificial neural network models. *Atmos Environ* 191:205–213. <https://doi.org/10.1016/j.atmosenv.2018.07.058>
- Almetwally AA, Bin-Jumah M, Allam AA (2020) Ambient air pollution and its influence on human health and welfare: an overview. *Environ Sci Pollut Res* 27(20):24815–24830. <https://doi.org/10.1007/s11356-020-09042-2>
- Awan FM, Minerva R, Crespi N (2020) Improving Road Traffic Forecasting Using Air Pollution and Atmospheric Data: Experiments Based on LSTM Recurrent Neural Networks. *Sensors* 20(13):3749. <https://doi.org/10.3390/s20133749>
- Baldasano JM (2020) COVID-19 lockdown effects on air quality by NO₂ in the cities of Barcelona and Madrid (Spain). *Sci Total Environ* 741:140353. <https://doi.org/10.1016/j.scitotenv.2020.140353>
- Barua S, Nath SD (2021) The impact of COVID-19 on air pollution: Evidence from global data. *J Clean Prod* 298. <https://doi.org/10.1016/j.jclepro.2021.126755>
- Bauer SE, Im U, Mezuman K, Gao CY (2019) Desert Dust, Industrialization, and Agricultural Fires: Health Impacts of Outdoor Air Pollution in Africa. *J Geophys Res: Atmospheres* 124(7):4104–4120. <https://doi.org/10.1029/2018jd029336>
- Bera B, Bhattacharjee S, Shit PK, Sengupta N, Saha S (2020) Significant impacts of COVID-19 lockdown on urban air pollution in Kolkata (India) and amelioration of environmental health. *Environ Dev Sustain* 23(5):6913–6940. <https://doi.org/10.1007/s10668-020-00898-5>
- Berman JD, Ebisu K (2020) Changes in U.S. air pollution during the COVID-19 pandemic. *Sci Total Environ* 739:139864. <https://doi.org/10.1016/j.scitotenv.2020.139864>
- Boznar M, Lesjak M, Mlakar P (1993) A neural network-based method for short-term predictions of ambient SO₂ concentrations in highly polluted industrial areas of complex terrain. *Atmos Environ Part B Urban Atmos* 27(2):221–230. [https://doi.org/10.1016/0957-1272\(93\)90007-s](https://doi.org/10.1016/0957-1272(93)90007-s)
- Cabaneros SM, Calautit JK, Hughes BR (2019) A review of artificial neural network models for ambient air pollution prediction. *Environ Model Softw* 119:285–304. <https://doi.org/10.1016/j.envsoft.2019.06.014>
- Campagnolo ML, Sun Q, Liu Y, Schaaf C, Wang Z, Román MO (2016) Estimating the effective spatial resolution of the operational BRDF, albedo, and nadir reflectance products from MODIS and VIIRS. *Remote Sens Environ* 175:52–64. <https://doi.org/10.1016/j.rse.2015.12.033>
- Chang Z, Zhang Y, Chen W (2018) Effective Adam-Optimized LSTM Neural Network for Electricity Price Forecasting. Paper presented at the 2018 IEEE 9th International Conference on Software Engineering and Service Science (ICSESS)
- Czech R, Zabochnicka-Świątek M, MK Ś (2020) Air pollution as a result of the development of motorization. *Global NEST Journal*. <https://doi.org/10.30955/gnj.003021>
- Dong Y, Zhang L, Liu Z, Wang J (2019) Integrated forecasting method for wind energy management: A case study in China. *Processes* 8(1):35
- Dong Y, Li J, Liu Z, Niu X, Wang J (2022) Ensemble wind speed forecasting system based on optimal model adaptive selection strategy: Case study in China. *Sustain Energy Technol Assess* 53:102535

- Duncan JMA, Boruff B, Saunders A, Sun Q, Hurley J, Amati M (2019) Turning down the heat: An enhanced understanding of the relationship between urban vegetation and surface temperature at the city scale. *Sci Total Environ* 656:118–128. <https://doi.org/10.1016/j.scitotenv.2018.11.223>
- Effendi A, Budianto B, Immanuel GS, Rakhman A, Kinasih SAKW, Boer R (2021) Coverage Sensitivity of High-Rise Tower NIES Monitoring System. IOP Conference Series: Earth Environ Sci 893(1):012072. <https://doi.org/10.1088/1755-1315/893/1/012072>
- Grainger C, Schreiber A (2019) Discrimination in Ambient Air Pollution Monitoring? AEA Papers Proc 109:277–282. <https://doi.org/10.1257/pandp.20191063>
- Guo Q, He Z (2021) Prediction of the confirmed cases and deaths of global COVID-19 using artificial intelligence. *Environ Sci Pollut Res* 28:11672–11682
- Guo Q, He Z, Li S, Li X, Meng J, Hou Z, Chen Y (2020) Air pollution forecasting using artificial and wavelet neural networks with meteorological conditions. *Aerosol and Air Qual Res* 20(6):1429–1439
- Guo Q, Wang Z, He Z, Li X, Meng J, Hou Z, Yang J (2021) Changes in air quality from the COVID to the post-COVID era in the beijing-tianjin-tangshan region in China. *Aerosol Air Qual Res* 21(12):210270
- Guo Q, He Z, Wang Z (2023) Predicting of daily PM2.5 concentration employing wavelet artificial neural networks based on meteorological elements in Shanghai China. *Toxics* 11(1):51
- Guo Q, He Z, Wang Z (2023b) Prediction of Hourly PM2.5 and PM10 Concentrations in Chongqing City in China Based on Artificial Neural Network. *Aerosol and Air Qual Res* 23:220448
- He Z, Guo Q, Wang Z, Li X (2022) Prediction of monthly PM2.5 concentration in Liaocheng in China employing artificial neural network. *Atmosphere* 13(8):1221
- Hojamberdiev M, Piccirillo C, Cai Y, Kadirova ZC, Yubuta K, Ruzimuradov O (2019) ZnS-containing industrial waste: Antibacterial activity and effects of thermal treatment temperature and atmosphere on photocatalytic activity. *J Alloy Compd* 791:971–982
- Hu Y, Rein G (2022) Development of gas signatures of smouldering peat wildfire from emission factors. *Int J Wildland Fire* 31(11):1014–1032. <https://doi.org/10.1071/wf21093>
- Jelonek Z, Drobniak A, Mastalerz M, Jelonek I (2020) Environmental implications of the quality of charcoal briquettes and lump charcoal used for grilling. *Sci Total Environ* 747. <https://doi.org/10.1016/j.scitotenv.2020.141267>
- Jiang S-Y, Ma A, Ramachandran S (2018) Negative Air Ions and Their Effects on Human Health and Air Quality Improvement. *Int J Mole Sci* 19(10):2966. <https://doi.org/10.3390/ijms19102966>
- Jiang P, Liu Z, Wang J, Zhang L (2022a) Decomposition-selection-ensemble prediction system for short-term wind speed forecasting. *Electric Power Systems Research* 211:108186
- Jiang P, Liu Z, Zhang L, Wang J (2022b) Advanced traffic congestion early warning system based on traffic flow forecasting and extenics evaluation. *Appl Soft Comput* 118:108544
- Kabiraj S (2021) Spatial expansion of industrial area and its impact on environmental indicators—a case study in West Bengal. *Arab J Geosci* 14(12):1159
- Kalajdzieski J, Zdravevski E, Corizzo R, Lameski P, Kalajdziski S, Pires IM, Garcia NM, Trajkovic V (2020) Air Pollution Prediction with Multi-Modal Data and Deep Neural Networks. *Remote Sensing* 12(24):4142. <https://doi.org/10.3390/rs12244142>
- Kazemi Z, Jonidi Jafari A, Farzadkia M, Kazemnezhad Leyli E, Shahsavani A, Kermani M (2022) Assessment of the risk of exposure to Air pollutants and identifying the affecting factors on making pollution by PCA, CFA. *Int J Environ Anal Chem* 1–20. <https://doi.org/10.1080/03067319.2022.2059364>
- Kwak MJ, Lee JK, Park S, Lim YJ, Kim H, Kim KN, Je SM, Park CR, Woo SY (2020) Evaluation of the Importance of Some East Asian Tree Species for Refinement of Air Quality by Estimating Air Pollution Tolerance Index, Anticipated Performance Index, and Air Pollutant Uptake. *Sustainability* 12(7):3067. <https://doi.org/10.3390/su12073067>
- Laib O, Khadir MT, Mihaylova L (2019) Toward efficient energy systems based on natural gas consumption prediction with LSTM Recurrent Neural Networks. *Energy* 177:530–542. <https://doi.org/10.1016/j.energy.2019.04.075>
- Lee S, Kim M, Lee KT, Irvine JT, Shin TH (2021) Enhancing electrochemical CO2 reduction using Ce (Mn, Fe) O2 with La (Sr) Cr (Mn) O3 cathode for high-temperature solid oxide electrolysis cells. *Adv Energy Mater* 11(24):2100339
- Li J, Han X, Zhang X, Sheveleva AM, Cheng Y, Tuna F, McInnes EJJ, McCormick McPherson LJ, Teat SJ, Daemen LL, Ramirez-Cuesta AJ, Schröder M, Yang S (2019a) Capture of nitrogen dioxide and conversion to nitric acid in a porous metal–organic framework. *Nat Chem* 11(12):1085–1090. <https://doi.org/10.1038/s41557-019-0356-0>
- Li T, Li Y, An D, Han Y, Xu S, Lu Z, Crittenden J (2019b) Mining of the association rules between industrialization level and air quality to inform high-quality development in China. *J Environ Manage* 246:564–574. <https://doi.org/10.1016/j.jenvman.2019.06.022>
- Liu Z, Li P, Wei D, Wang J, Zhang L, Niu X (2023) Forecasting system with sub-model selection strategy for photovoltaic power output forecasting. *Earth Sci Inf* 16(1):287–313
- Liu DR, Lee SJ, Huang Y, Chiu CJ (2019) Air pollution forecasting based on attention-based LSTM neural network and ensemble learning. *Expert Systems* 37 (3). <https://doi.org/10.1111/exsy.12511>
- Luong LMT, Phung D, Dang TN, Sly PD, Morawska L, Thai PK (2018) Seasonal association between ambient ozone and hospital admission for respiratory diseases in Hanoi Vietnam. *Plos One* 13(9):e0203751. <https://doi.org/10.1371/journal.pone.0203751>
- Maleki H, Sorooshian A, Goudarzi G, Baboli Z, Tahmasebi Birgani Y, Rahmati M (2019) Air pollution prediction by using an artificial neural network model. *Clean Technol Environ Policy* 21(6):1341–1352. <https://doi.org/10.1007/s10098-019-01709-w>
- Mirsanjari MM, Zarandian A, Mohammadyari F, Visockiene JS (2020) Investigation of the impacts of urban vegetation loss on the ecosystem service of air pollution mitigation in Karaj metropolis, Iran. *Environ Monit Assess* 192 (8). <https://doi.org/10.1007/s10661-020-08399-8>
- Mohammadi B, Ahmadi F, Mehdizadeh S, Guan Y, Pham QB, Linh NTT, Tri DQ (2020) Developing novel robust models to improve the accuracy of daily streamflow modeling. *Water Resour Manage* 34:3387–3409
- Mor S, Ravindra K (2023) Municipal solid waste landfills in lower- and middle-income countries: Environmental impacts, challenges and sustainable management practices. *Process Saf Environ Prot* 174:510–530. <https://doi.org/10.1016/j.psep.2023.04.014>
- Müller A, Österlund H, Marsalek J, Viklander M (2020) The pollution conveyed by urban runoff: A review of sources. *Sci Total Environ* 709. <https://doi.org/10.1016/j.scitotenv.2019.136125>
- Muthukumar P, Cocom E, Nagrecha K, Comer D, Burga I, Taub J, Calvert CF, Holm J, Pourhomayoun M (2021) Predicting PM2.5 atmospheric air pollution using deep learning with meteorological data and ground-based observations and remote-sensing satellite big data. *Air Qual, Atmos Health* 15(7):1221–1234. <https://doi.org/10.1007/s11869-021-01126-3>
- Nanda, S. K., Tripathy, D. P., Mohapatra, R., & Ray, N. K. (2021, December). Application of 1-Dimensional Convolution Neural Network based Machine Learning Approach for Prediction of Air Quality Index. In 2021 19th OITS International Conference on Information Technology (OCIT) (pp. 341–346). IEEE

- Nho R (2020) Pathological effects of nano-sized particles on the respiratory system. *Nanomed: Nanotechnol, Biol Med* 29:102242. <https://doi.org/10.1016/j.nano.2020.102242>
- Nichol JE, Bilal M, Ali MA, Qiu Z (2020) Air Pollution Scenario over China during COVID-19. *Remote Sensing* 12(13):2100. <https://doi.org/10.3390/rs12132100>
- Othman M, Latif MT (2021) Air pollution impacts from COVID-19 pandemic control strategies in Malaysia. *J Clean Prod* 291. <https://doi.org/10.1016/j.jclepro.2021.125992>
- Oves M, Zain Khan M, M.I. Ismail I (2018) Modern Age Environmental Problems and their Remediation. <https://doi.org/10.1007/978-3-319-64501-8>
- Pal S, Ghosh S, Nag A (2018) Sentiment Analysis in the Light of LSTM Recurrent Neural Networks. *Int J Synth Emotions* 9(1):33–39. <https://doi.org/10.4018/ijse.2018010103>
- Park HS, Kang D, Kang JH, Kim K, Kim J, Song H (2021) Selective sulfur dioxide absorption from simulated flue gas using various aqueous alkali solutions in a polypropylene hollow fiber membrane contactor: removal efficiency and use of sulfur dioxide. *Int J Environ Res Public Health* 18(2):597. <https://doi.org/10.3390/ijerph18020597>
- Pedruzzi R, Baek BH, Henderson BH, Aravanis N, Pinto JA, Araujo IB, Nascimento EGS, Reis Junior NC, Moreira DM, de Almeida Albuquerque TT (2019) Performance evaluation of a photochemical model using different boundary conditions over the urban and industrialized metropolitan area of Vitória Brazil. *Environ Sci Pollut Res* 26(16):16125–16144. <https://doi.org/10.1007/s11356-019-04953-1>
- Polii B, Najooan J, Ogie T (2021) Analysis of Greenhouse Gases and Odor Levels in the Sumompo TPA, Manado City, North Sulawesi. *Agri-Sosioekonomi* 17 (1). <https://doi.org/10.35791/agrrosek.17.1.2021.32230>
- Prockop LD, Chichkova RI (2007) Carbon monoxide intoxication: An updated review. *J Neurol Sci* 262(1–2):122–130. <https://doi.org/10.1016/j.jns.2007.06.037>
- Rahaman ZA, Kafy AA, Saha M, Rahim AA, Almulhim AI, Rahaman SN, Fattah MA, Rahman MT, S K, Faisal A-A, Al Rakib A (2022) Assessing the impacts of vegetation cover loss on surface temperature, urban heat island and carbon emission in Penang city, Malaysia. *Build Environ* 222. <https://doi.org/10.1016/j.buildenv.2022.109335>
- Reddington CL, Conibear L, Robinson S, Knote C, Arnold SR, Spracklen DV (2021) Air Pollution From Forest and Vegetation Fires in Southeast Asia Disproportionately Impacts the Poor. *GeoHealth* 5 (9). <https://doi.org/10.1029/2021gh000418>
- Roy P, Pal SC, Chakraborty R, Chowdhuri I, Saha A, Shit M (2022) Climate change and groundwater overdraft impacts on agricultural drought in India: Vulnerability assessment, food security measures and policy recommendation. *Sci Total Environ* 849. <https://doi.org/10.1016/j.scitotenv.2022.157850>
- Saevarsdotir G, Kvande H, Welch BJ (2019) Aluminum Production in the Times of Climate Change: The Global Challenge to Reduce the Carbon Footprint and Prevent Carbon Leakage. *Jom* 72(1):296–308. <https://doi.org/10.1007/s11837-019-03918-6>
- Santamouris M, Osmond P (2020) Increasing Green Infrastructure in Cities: Impact on Ambient Temperature, Air Quality and Heat-Related Mortality and Morbidity. *Buildings* 10(12):233. <https://doi.org/10.3390/buildings10120233>
- Selvam S, Muthukumar P, Venkatramanan S, Roy PD, Manikanda Bharath K, Jesuraja K (2020) SARS-CoV-2 pandemic lockdown: Effects on air quality in the industrialized Gujarat state of India. *Sci Total Environ* 737. <https://doi.org/10.1016/j.scitotenv.2020.140391>
- Shen Y, Mercatoris B, Cao Z, Kwan P, Guo L, Yao H, Cheng Q (2022) Improving Wheat Yield Prediction Accuracy Using LSTM-RF Framework Based on UAV Thermal Infrared and Multispectral Imagery. *Agriculture* 12(6):892. <https://doi.org/10.3390/agriculture12060892>
- Shi Z, Song C, Liu B, Lu G, Xu J, Van Vu T, Elliott RJR, Li W, Bloss WJ, Harrison RM (2021) Abrupt but smaller than expected changes in surface air quality attributable to COVID-19 lockdowns. *Sci Adv* 7 (3). <https://doi.org/10.1126/sciadv.abd6696>
- Soeder DJ (2021) Greenhouse gas sources and mitigation strategies from a geosciences perspective. *Adv Geo-Energy Res* 5(3):274–285. <https://doi.org/10.46690/ager.2021.03.04>
- Souza JBC, de Almeida SLH, Freire de Oliveira M, Santos AFd, Filho ALdB, Meneses MD, Silva RPd (2022) Integrating Satellite and UAV Data to Predict Peanut Maturity upon Artificial Neural Networks. *Agronomy* 12(7):1512. <https://doi.org/10.3390/agronomy12071512>
- Tiwari S, Thomas A, Rao P, Chate DM, Soni VK, Singh S, Ghude SD, Singh D, Hopke PK (2018) Pollution concentrations in Delhi India during winter 2015–16: A case study of an odd-even vehicle strategy. *Atmos Pollut Res* 9(6):1137–1145. <https://doi.org/10.1016/j.apr.2018.04.008>
- Tiwari A, Kumar P (2020) Integrated dispersion-deposition modeling for air pollutant reduction via green infrastructure at an urban scale. *Sci Total Environ* 723. <https://doi.org/10.1016/j.scitotenv.2020.138078>
- Tondelli S, Farhadi E, Akbari Monfared B, Ataiean M, Tahmasebi-Moghaddam H, Dettori M, Saganeiti L, Murgante B (2022) Air Quality and Environmental Effects Due to COVID-19 in Tehran, Iran: Lessons for Sustainability. *Sustainability* 14(22):15038. <https://doi.org/10.3390/su142215038>
- Vadrevu K, Lasko K (2018) Intercomparison of MODIS AQUA and VIIRS I-Band Fires and Emissions in an Agricultural Landscape—Implications for Air Pollution Research. *Remote Sensing* 10(7):978. <https://doi.org/10.3390/rs10070978>
- Vidal A, Kristjanpoller W (2020) Gold volatility prediction using a CNN-LSTM approach. *Expert Syst App* 157. <https://doi.org/10.1016/j.eswa.2020.113481>
- Wang W, Liang S, Meyers T (2008) Validating MODIS land surface temperature products using long-term nighttime ground measurements. *Remote Sens Environ* 112(3):623–635. <https://doi.org/10.1016/j.rse.2007.05.024>
- Wang J, Zhang L, Liu Z, Niu X (2022) A novel decomposition-ensemble forecasting system for dynamic dispatching of smart grid with sub-model selection and intelligent optimization. *Expert Syst Appl* 201:117201
- Wang Z, Wei W (2020) Effects of modifying industrial plant configuration on reducing air pollution-induced agricultural loss. *J Clean Prod* 277. <https://doi.org/10.1016/j.jclepro.2020.124046>
- Wang G, Xia X, Liu S, Zhang L, Zhang S, Wang J, Xi N, Zhang Q (2021a) Intense methane ebullition from urban inland waters and its significant contribution to greenhouse gas emissions. *Water Research* 189. <https://doi.org/10.1016/j.watres.2020.116654>
- Wang J, Ye J, Zhang Q, Zhao J, Wu Y, Li J, Liu D, Li W, Zhang Y, Wu C, Xie C, Qin Y, Lei Y, Huang X, Guo J, Liu P, Fu P, Li Y, Lee HC, Choi H, Zhang J, Liao H, Chen M, Sun Y, Ge X, Martin ST, Jacob DJ (2021b) Aqueous production of secondary organic aerosol from fossil-fuel emissions in winter Beijing haze. *Proc Nat Acad Sci* 118(8). <https://doi.org/10.1073/pnas.2022179118>
- Wen C, Liu S, Yao X, Peng L, Li X, Hu Y, Chi T (2019) A novel spatiotemporal convolutional long short-term neural network for air pollution prediction. *Sci Total Environ* 654:1091–1099. <https://doi.org/10.1016/j.scitotenv.2018.11.086>
- Yadav A, Jha CK, Sharan A (2020) Optimizing LSTM for time series prediction in Indian stock market. *Procedia Comput Sci* 167:2091–2100. <https://doi.org/10.1016/j.procs.2020.03.257>

- Yao R, Wang L, Huang X, Gong W, Xia X (2019) Greening in Rural Areas Increases the Surface Urban Heat Island Intensity. *Geophys Res Lett* 46(4):2204–2212. <https://doi.org/10.1029/2018gl081816>
- Yu Y, Wang J, Liu Z, Zhao W (2021). A combined forecasting strategy for the improvement of operational efficiency in wind farm. *J Renew Sustain Energy* 13(6)
- Yuan M, Song Y, Huang Y, Shen H, Li T (2019) Exploring the association between the built environment and remotely sensed PM2.5 concentrations in urban areas. *J Clean Prod* 220:1014–1023. <https://doi.org/10.1016/j.jclepro.2019.02.236>
- Zahra SI, Iqbal MJ, Ashraf S, Aslam A, Ibrahim M, Yamin M, Vithanage M (2022) Comparison of Ambient Air Quality among Industrial and Residential Areas of a Typical South Asian City. *Atmosphere* 13(8):1168. <https://doi.org/10.3390/atmos13081168>
- Zeng J, He Q (2019) Does industrial air pollution drive health care expenditures? Spatial evidence from China. *J Clean Prod* 218:400–408. <https://doi.org/10.1016/j.jclepro.2019.01.288>
- Zhang L, Wang J, Liu Z (2023) Power grid operation optimization and forecasting using a combined forecasting system. *J Forecast* 42(1):124–153
- Zhang F, Li Y, Li Y, Xu Y, Chen J (2022) Nexus among air pollution, enterprise development and regional industrial structure upgrading: A China's country panel analysis based on satellite retrieved data. *J Clean Prod* 335. <https://doi.org/10.1016/j.jclepro.2021.130328>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.